The SED Virtual Machine Environment: What is it and What Does it Do?

11/1/2011

Catherine Corlan

Some Initial Concepts

- A Virtual Machine Environment is
  - A “hypervisor”, running on physical server(s), which maintains multiple VMs
    - Customers effectively get their own machine
- Characteristics of a Good Candidate for Virtualization
  - Uses < 30% of Server Resources (CPU, Memory, I/O)
  - OR Usage is strongly “bursty” – higher utilization for 5 - 15% of the time
- Poor Candidates for Virtualization can still be good candidates for larger Cloud environments

VME Advantages

- High availability: separating the HW/SW layers
  - HW can be reset, repaired or upgraded with no VM downtime, provided the facility has the spare cycles
- Multiple, Specialized VMs instead of 1 machine running all processes
  - Improve Work Processes
    - Even small projects can now have multiple machines, separating Development & Test from Production
  - Improve Security Posture
    - E.g., separate DB VM, mounted read-only in the public zone: if you get hacked, your data is still safe
- Reduced Paperwork
  - We handle your security plan and business case
  - We help you meet Code 700 requirements – NAMS registration, STRAW Registration, Port 80 waivers, etc.
- Institutionally, we save on HW costs and reduce our raised-floor footprint (lower electricity, cooling)

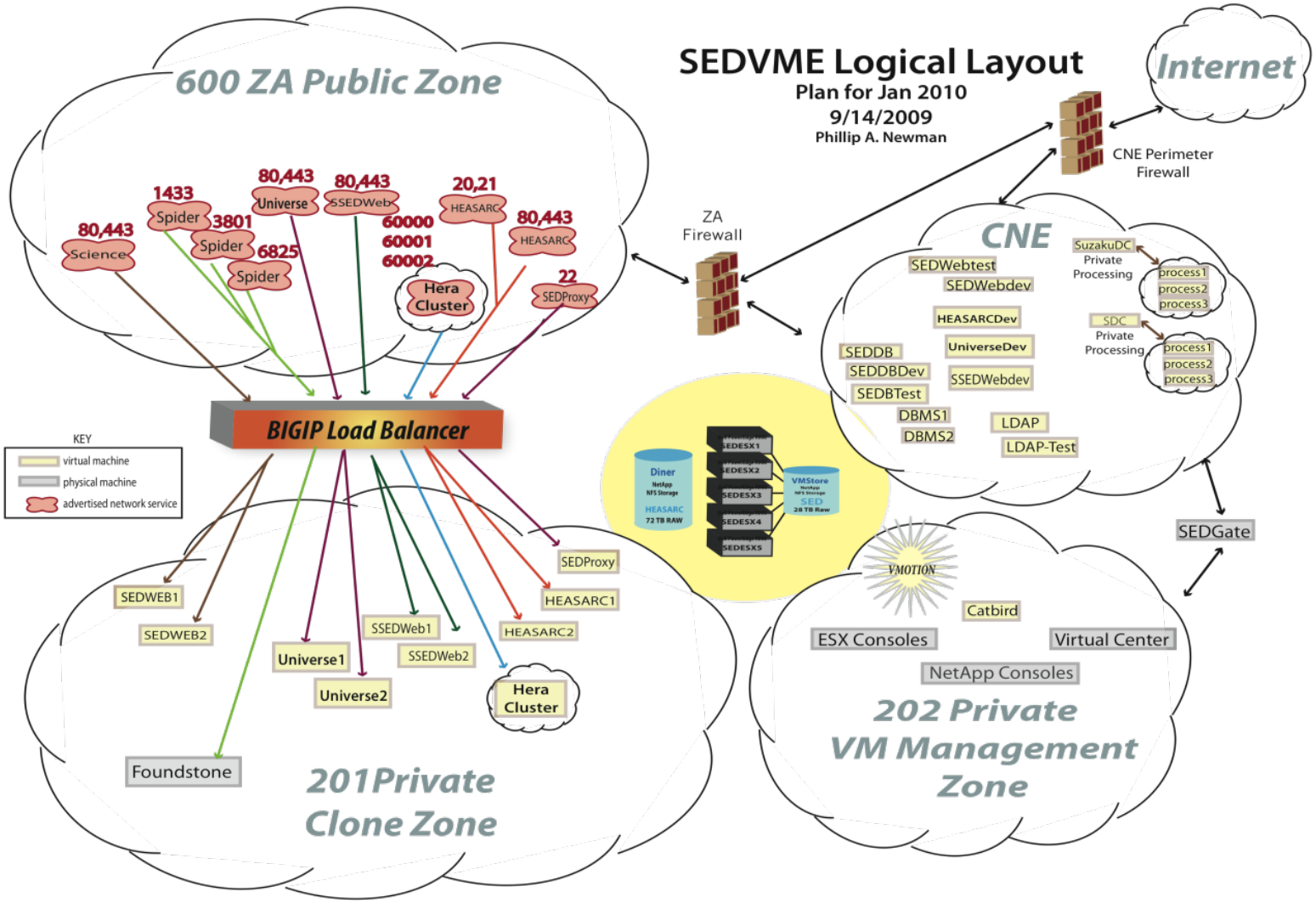
SED VME Purpose

- Provide Operational Environment for 660 & 690
  - Solve the computer space problem associated with moving to Bldg 34, losing Bldg 2
- Extra capacity to enable the Directorate to
  - Consolidate HW servers
  - Prototype ‘Cloud’ computing
    - Infrastructure as a Service (bare machine)
    - **Platform as a Service** (machine + installed OS)
      - PaaS plus SA support for installation & configuration
    - Software as a Service (provide, e.g., a server running MatLab for a large number of users)
      - Variation: adding users to Division-level Web servers, DB servers, etc.
  - Work out policies & procedures for shared environment with maximal customer control

VME Design: Infrastructure

- Hardware
  - 5 Dell ESX servers, each with 16 CPUs running at 2.9GHz, 128 GB Memory
  - NetApp SATA RAID, 14TB, running SnapMirror to B32 back-up
  - Big IP to provide security & load balancing on public-facing VMs
  - Network: On CNE (10 Gb uplink to core), uses the Public Zone
  - Planning to add SAN & SAS Disk this year
- Software
  - VMware hypervisor: Resource Scheduler, vMotion, High Availability
  - Red Hat Satellite Server (automatically pushes SW updates)
  - Red Hat Linux: infinite licenses per machine; our preferred OS
    - Also running SUSE and Windows VMs

VME Diagram


SEDVME Logical Layout
Plan for Jan 2010
9/14/2009
Phillip A. Newman
600 ZA Public Zone
80,443
Science
1433
Spider
3801
Spider
6825
Spider
80,443
Universe
80,443
SSEDWeb
60000
60001
60002
Hera Cluster
20,21
HEASARC
80,443
HEASARC
22
SEDProxy
ZA Firewall
Internet
CNE Perimeter Firewall
CNE
SEDWebtest
SEDWebdev
HEASARCDev
UniverseDev
SSEDWebdev
SEDDB
SEDDBDev
SEDBTest
DBMS1
DBMS2
LDAP
LDAP-Test
SuzakuDC
Private Processing
process1
process2
process3
SDC
Private Processing
process1
process2
process3
BIGIP Load Balancer
KEY
virtual machine
physical machine
advertised network service
Diner
HEASARC
72 TB RAW
SEDESX1
SEDESX2
SEDESX3
SEDESX4
SEDESX5
VMStore
SED
20 TB Raw
SEDGate
VMOTION
Catbird
ESX Consoles
Virtual Center
NetApp Consoles
202 Private VM Management Zone
SEDWEB1
SEDWEB2
SEDProxy
HEASARC1
HEASARC2
SSEDWeb1
SSEDWeb2
Universe1
Universe2
Hera Cluster
Foundstone
201Private Clone Zone

Current Status

- Successful, stable operational environment with good performance
  - Running 93 VMs, ~40% Capacity; Looking for more customers
- 99.97% Uptime (1 malicious shut-down, 6 hrs)
- Preliminary performance benchmarks show no significant difference in performance between our VM and a comparable dedicated system in the NCCS
- Customers include HEASARC, Swift, SED Website
- FISMA Low; cooperating with new FISMA Moderate VME in B32/C101 for MOCs

Lessons Learned

- It Works, and saves on equipment costs & utilities.
- Multiple types of Disk are necessary
- High Performance depends on configuration details
  - OS Kernel tweaks, VM Memory handling
  - This argues for limiting OS flavors when there are only a few physical machines
- High Availability requires lots of idle cycles
  - Looking at 2-tier pricing, discount for risk of less availability
- VMware's vMotion requires same instruction set CPUs to function – limits refresh options
- SA Support Levels: if the machines are not clones, SA support *does not* scale
  - Effort ~ #VMs => Division-level special-purpose VMs to reduce proliferation & control costs

Customer Service

- Migration Process
  - Initial meeting to understand the customer requirements, design VM solution
    - Take advantage of virtualization to improve their workflow
    - Encourage VMs with the minimum resources needed (CPU, memory) to accomplish the job effectively
      - VMware locks up the allocated resources whether or not they are used
      - Largest VM is 4 CPUs, 8 GB mem; start with 1 CPU, 1 GB mem
  - 1 work-week to set up & configure machines for initial log-in
    - Additional time if NAMS, Port 80 Waivers, etc. are required
- 8 X 5 Customer Support (unless there are extraordinary circumstances)
  - Phone & email contact (gsfc-sedvmehelp@lists.nasa.gov); use issue tracking
  - SAs from the Divisions served allows support to be culturally customized

Cost Model

- In Development
- Designed to cover expenses
- Bill proportional to usage, but bill in advance => Usage Levels (low, med, high), with potential for quarterly adjustment if the usage level is badly wrong.
- Charge =

  (CPU cost/GHz * GHz Usage Level)

  + (Disk Cost/GB * Disk Allocation)

  + (Operational Expenses per VM * #VMs)

  + (SA Cost * SA Usage Level)
- Divisions pay for multi-project VMs, i.e., Division-wide Web Server, MySQL Server, ftp Server…
- Issues: Project lifecycle budgets, making multi-VMs affordable

Outlook & Plans

- Adding Disk to increase our flexibility
- Intend to develop Amazon-like web interface to let customers know pricing, availability in advance
  - Problems: customers don't know to separate their processes into multiple VMs; customers want more resources than they need
- Currently looking for customers; adding about 1/week
- Considering SaaS for MatLab, IDL
- Looking at scalability, performance: lessons learned to help get Local Cloud right